

AD-A031 952

WISCONSIN UNIV MADISON MATHEMATICS RESEARCH CENTER

F/G 12/2

AN IMPLICIT-FUNCTION THEOREM FOR GENERALIZED VARIATIONAL INEQUA--ETC(U)

SEP 76 S M ROBINSON

DAA629-75-C-0024

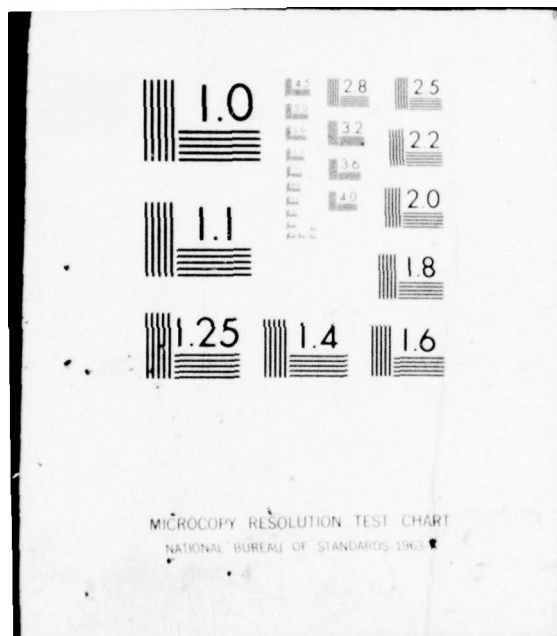
UNCLASSIFIED

MRC-TSR-1672

NL

1 OF 1
AD
A031952





ADA031952

MRC Technical Summary Report #1672

AN IMPLICIT-FUNCTION THEOREM
FOR GENERALIZED VARIATIONAL
INEQUALITIES

Stephen M. Robinson

Mathematics Research Center
University of Wisconsin-Madison
610 Walnut Street
Madison, Wisconsin 53706

September 1976

(Received August 4, 1976)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

and

National Science Foundation
Washington, D. C.
20550

DDC
RECEIVED
NOV 12 1976
See Form 1473

MATHEMATICS RESEARCH CENTER
AN IMPLICIT-FUNCTION THEOREM FOR GENERALIZED
VARIATIONAL INEQUALITIES

Stephen M. Robinson*

Technical Summary Report #1672
September 1976

ABSTRACT

✓ An implicit-function theorem is established for a multifunction consisting of the sum of a differentiable function and a maximal monotone operator. Applications to nonlinear complementarity problems, mathematical programming problems, and economic equilibria are pointed out. An application to the analysis of a general Newton method for solving variational inequalities is treated in some detail. ↗

AMS (MOS) Subject Classifications: 47H05, 90A15, 90C30

Key Words: Variational inequalities
Monotone operators
Nonlinear complementarity problem
Nonlinear programming
Economic equilibria

Work Unit Number 5 (Mathematical Programming and Operations Research)

* Sponsored in part by the United States Army under Contract No. DAAG29-75-C-0024, and in part by the National Science Foundation under Grant No. DCR74-20584.

— a —

RECEIVED	DATE	BY
SEP 10 1976	SEP 10 1976	STEPHEN M. ROBINSON
LIBRARY	DATE	BY
SEP 10 1976	SEP 10 1976	STEPHEN M. ROBINSON
RECEIVED	DATE	BY
SEP 10 1976	SEP 10 1976	STEPHEN M. ROBINSON

AN IMPLICIT-FUNCTION THEOREM FOR GENERALIZED
VARIATIONAL INEQUALITIES

Stephen M. Robinson^{*}

1. Introduction. In this paper we shall study the behavior of solutions of the inclusion

$$0 \in f(x) + T(x), \quad (1.1)$$

where f is a continuously Fréchet differentiable function from an open set $N \subset \mathbb{R}^n$ into \mathbb{R}^n and T is a maximal monotone operator from \mathbb{R}^n into itself (recall that an operator T is monotone if for each $(x_1, w_1), (x_2, w_2)$ in graph T one has

$$\langle x_1 - x_2, w_1 - w_2 \rangle \geq 0,$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product, and maximal monotone if its graph is not properly contained in that of any other monotone operator).

We shall be particularly interested in conditions which, when imposed on f and T , will ensure that the set of solutions to (1.1) remains nonempty and is well behaved (in a sense to be defined) when f is subjected to small perturbations. To introduce these perturbations, we shall make use of a topological space P and a function $f : P \times N \rightarrow \mathbb{R}^n$, so that we can replace (1.1) by

$$0 \in f(p, x) + T(x), \quad (1.2)$$

^{*} Sponsored in part by the United States Army under Contract No. DAAG29-75-C-0024, and in part by the National Science Foundation under Grant No. DCR74-20584.

and study the set of x which solve (1.2) as p varies near a base value p_0 .

A particular case of (1.2) of special interest for applications is that in which T is taken to be the operator $\partial\psi_C$, where for a closed convex set $C \subset \mathbb{R}^n$ one defines the indicator function ψ_C of C by

$$\psi_C(x) := \begin{cases} 0, & x \in C \\ +\infty, & x \notin C, \end{cases}$$

and where ∂ denotes the subdifferential operator [16, §23]. This yields the so-called variational inequality

$$0 \in f(p, x) + \partial\psi_C(x), \quad (1.3)$$

which expresses analytically the geometric idea that $f(p, x)$ is an inward normal to C at x .

Many problems from mathematical programming, complementarity, mathematical economics and other fields can be represented in the form (1.3): for example, the nonlinear complementarity problem

$$\begin{aligned} F(x) &\in K^* \\ x &\in K \\ \langle x, F(x) \rangle &= 0 \end{aligned} \quad (1.4)$$

where $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$, K is a nonempty polyhedral convex cone in \mathbb{R}^n , and $K^* := \{y \in \mathbb{R}^n \mid \langle y, k \rangle \geq 0 \ \forall k \in K\}$, can be written as

$$0 \in F(x) + \partial\psi_K(x).$$

Further information on nonlinear complementarity problems (often with $K = \mathbb{R}_+^n$, the non-negative orthant) may be found in, e.g., [2], [3], [8], [9]. The Kuhn-Tucker necessary conditions for mathematical programming [7] form a special case of (1.4); e.g., for the problem

$$\begin{aligned} & \text{minimize} && \theta(y) \\ & \text{subject to} && g(y) \leq 0 \\ & && h(y) = 0 \end{aligned} \tag{1.5}$$

where θ , g and h are differentiable function from \mathbb{R}^m into \mathbb{R} , \mathbb{R}^q and \mathbb{R}^r respectively, one has the Kuhn-Tucker conditions

$$\begin{aligned} \theta'(y) + ug'(y) + vh'(y) &= 0 \\ g(y) &\leq 0 \\ h(y) &= 0 \\ u &\geq 0 \\ \langle u, g(y) \rangle &= 0, \end{aligned}$$

and these can be written in the form (1.4) by taking $n = m + q + r$,

$K = \mathbb{R}^m \times \mathbb{R}_+^q \times \mathbb{R}^r$, $x = (y, u, v)$ and

$$F(x) = \begin{bmatrix} [\theta'(y) + ug'(y) + vh'(y)]^T \\ -g(y) \\ -h(y) \end{bmatrix}.$$

There are also important applications of (1.3) to economic equilibrium problems [5], [18], [19], among others. It is of interest to note that in most of the applications mentioned one finds that C is a polyhedral

convex set, and we shall see that particularly strong results can be obtained for such problems.

The organization of this paper is as follows: in the next section we shall state an implicit-function theorem (Theorem 1) for systems of the form (1.2); we then specialize this result to the case in which T has a certain polyhedrality property and obtain a form (Theorem 2) which is useful in applications. In the remainder of Section 2 we shall give the proof of Theorem 1. In Section 3 we shall prove Theorem 2, after having established some necessary results about polyhedral sets and functions. Finally, in Section 4 we shall apply these results to prove implementability and quadratic convergence of a general Newton algorithm for computational solution of variational inequalities.

2. Main results. Before stating the main theorem, we require a preliminary definition dealing with a certain continuity property of multi-valued functions (or multifunctions, as we shall call them).

DEFINITION: Let X and Y be normed linear spaces. A multi-function $F : X \rightarrow Y$ is upper Lipschitzian with modulus λ , or $U.L.(\lambda)$, at a point $x_0 \in X$ with respect to a set $V \subset X$, if for each $v \in V$ one has

$$F(v) \subset F(x_0) + \lambda \|v - x_0\| B_Y,$$

where B_Y is the unit ball in Y . We say F is locally $U.L.(\lambda)$ at x_0 if it is $U.L.(\lambda)$ at x_0 with respect to some neighborhood of x_0 .

This property is close to the Lipschitz continuity for multifunctions defined by Rockafellar [17, §3], except that we do not require $F(x_0)$ to be a singleton; in the problems we shall consider $F(x_0)$ will often be multivalued.

Our principal result is the following theorem, in which we use f_2 to denote the partial Fréchet derivative, with respect to the second argument, of a function of two variables; B denotes the unit ball in \mathbb{R}^n with respect to the Euclidean norm, which is used throughout the remainder of the paper.

THEOREM 1: Let P be a topological space, N an open set in \mathbb{R}^n and T a maximal monotone operator from \mathbb{R}^n into itself. Let f be a continuous function from $P \times N$ into \mathbb{R}^n such that f_2 is continuous on $P \times N$. Let $p_0 \in P$; suppose that there are a nonempty, bounded subset X_0 of N and numbers λ and $\eta > 0$ such that for each $x_0 \in X_0$,

i) $(Lf_{x_0} + T)^{-1}$ is U.L.(λ) at 0 with respect to ηB , where

$$Lf_{x_0}(\cdot) := f(p_0, x_0) + f_2(p_0, x_0)[(\cdot) - x_0],$$

ii) $(Lf_{x_0} + T)^{-1}(0) = X_0$, and

iii) $f_2(p_0, x_0)$ is positive semidefinite.

Then there exist a number $\delta > 0$ and a neighborhood $U(p_0)$ such that with

$$\Sigma(p) := \begin{cases} \{x \in X_0 + \delta B \mid 0 \in f(p, x) + T(x)\}, & p \in U \\ \emptyset, & p \notin U \end{cases}$$

one has:

1) Σ is upper semicontinuous from U to \mathbb{R}^n ,

2) $\Sigma(p_0) = X_0$,

and

3) For each $\varepsilon > 0$, for some neighborhood $U_\varepsilon(p_0)$ and for each $p \in U_\varepsilon$,

$$\emptyset \neq \Sigma(p) \subset \Sigma(p_0) + (\lambda + \varepsilon)\alpha_0(p)B,$$

where

$$\alpha_0(p) := \max\{\|f(p, x) - f(p_0, x)\| \mid x \in X_0\}.$$

Note that if P is actually a normed linear space and if $f(p, x)$ is Lipschitzian in p uniformly over $x \in X_0$, then for some constant μ and each $p \in U_\varepsilon$ we have

$$\Sigma(p) \subset \Sigma(p_0) + (\lambda + \varepsilon)\mu\|p - p_0\|B,$$

so that Σ is locally U.L. $[(\lambda + \varepsilon)\mu]$ at p_0 .

We next state a specialization of Theorem 1 to the case in which T has a certain additional property, described in the following definition.

DEFINITION 2: A multifunction $Q : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is polyhedral if its graph is the union of a finite (possibly empty) collection of polyhedral convex sets (called components).

Here we use "polyhedral convex set" as in [16, §19].

THEOREM 2: Let P and N be as in Theorem 1, and let $f : P \times N \rightarrow \mathbb{R}^n$ with f and f_2 continuous on $P \times N$. Let T be a polyhedral, maximal

monotone operator from \mathbb{R}^n into \mathbb{R}^n , and suppose that there are an $n \times n$ positive semidefinite matrix A and a point $a \in \mathbb{R}^n$ such that with

$$X_0 := \{x \in \mathbb{R}^n \mid 0 \in Ax + a + T(x)\}$$

one has

a. X_0 is a nonempty, bounded subset of N , and

b. For each $x_0 \in X_0$, $Lf_{x_0}(\cdot) \equiv A(\cdot) + a$.

Then the conclusions of Theorem 1 hold; the number λ can be taken to be a local upper Lipschitz modulus at 0 for $[A(\cdot) + a + T(\cdot)]^{-1}$, and this number exists without further assumptions.

We shall prove in Section 3 that if φ is a polyhedral convex function on \mathbb{R}^n (i.e., one whose epigraph is a polyhedral convex set) which never takes the value $-\infty$, then $\partial\varphi$ is a polyhedral multifunction in the sense of Definition 2. In this case it is well known that $\partial\varphi$ is also maximal monotone [16, Cor. 31.5.2], so the result in Theorem 2 is valid for $T = \partial\varphi$. In particular, Theorem 2 may be applied to the variational inequality (1.3), for the case in which C is a polyhedral convex set, by taking $\varphi = \psi_C$. One thus obtains an implicit-function theorem for (1.3); by specializing further to the case $P := \mathbb{R}^n$, $p_0 := 0$, and $f(p, x) := f(x) - p$, one obtains an "inverse-function" theorem dealing with inclusions of the type

$$p \in f(x) + \partial\psi_C(x),$$

for p near 0. This result, incidentally, provides a convenient tool

for verifying that the operator $(f + \partial_{\psi_C})^{-1}$ is "Lipschitz continuous at 0," in the sense required in [17, §3].

PROOF OF THEOREM 1: Denote by $\text{dom } G$ the effective domain of a multifunction G : i.e., the set $\{x \mid G(x) \neq \emptyset\}$. Choose any $x_0 \in X_0$; by (iii), $f_2(p_0, x_0)$ is positive semidefinite, so Lf_{x_0} is a maximal monotone operator. As T is also maximal monotone (hence nonempty) and as the effective domain of Lf_{x_0} is all of \mathbb{R}^n , we have from [15] that the operator $Q(x_0) := Lf_{x_0} + T$ is maximal monotone; hence so is its inverse. By (ii), $Q(x_0)^{-1}(0) = X_0$, a bounded set, and as by (i) $Q(x_0)^{-1}$ is locally U.L.(λ) at 0, we find that it is locally bounded there; in fact, it must be locally bounded at every point of $\text{int } \eta B$, since the image of some ball around such a point will be contained in the image of ηB , which in turn is contained in the bounded set $X_0 + \lambda \eta B$. But then from [14, Th. 1] we have that $\text{int } \eta B$ cannot contain any boundary point of $\text{dom } Q(x_0)^{-1}$; however, as $\text{int } \eta B$ meets $\text{dom } Q(x_0)^{-1}$ (at 0) and is connected we finally conclude that $\text{int } \eta B \subset \text{int dom } Q(x_0)^{-1}$. Thus, for each y with $\|y\| < \eta$ the set $Q(x_0)^{-1}(y)$ is nonempty and bounded; it is also closed and convex by maximal monotonicity, hence compact. In particular, X_0 is a compact convex set.

Define, for two subsets A and C of \mathbb{R}^n and a point $x \in \mathbb{R}^n$, $d[x, C] := \inf\{\|x - c\| \mid c \in C\}$ and $d[A, C] = \sup\{d[a, C] \mid a \in A\}$, where the supremum and infimum of \emptyset are defined to be $-\infty$ and $+\infty$ respectively. Denote by π the projection from \mathbb{R}^n onto X_0 ; π is well known to be nonexpansive, hence a fortiori continuous.

Using continuity and compactness, one can show that the function

$$\beta(\delta) := \max\{\|f_2(p_0, x) - f_2(p_0, \pi(x))\| \mid x \in X_0 + \delta B\}$$

is well defined for small δ , and is continuous at 0 with $\beta(0) = 0$.

Thus, we can choose a positive δ small enough that the set $X_\delta := X_0 + \delta B$

is contained in N , and such that $\lambda\beta(\delta) \leq \frac{1}{2}$ and $\delta\beta(\delta) \leq \frac{1}{2}\eta$. It is not

difficult to show that for this fixed δ the function

$$\alpha_\delta(p) := \max\{\|f(p, x) - f(p_0, x)\| \mid x \in X_\delta\}$$

is well defined for all $p \in P$, and is continuous at p_0 with $\alpha_\delta(p_0) = 0$.

Thus, we can choose a neighborhood $U(p_0)$ such that for each $p \in U$,

$\alpha_\delta(p) < \frac{1}{2}\eta$ and $\lambda\alpha_\delta(p) \leq \frac{1}{2}\delta$. Now choose any $p \in U$, and define a

multifunction F_p from X_δ into \mathbb{R}^n by

$$F_p(x) := Q(\pi(x))^{-1}[Lf_{\pi(x)}(x) - f(p, x)] .$$

If x is any point of X_δ , we have

$$\|Lf_{\pi(x)}(x) - f(p, x)\| \leq \|f(p, x) - f(p_0, x)\| + \|f(p_0, x) - Lf_{\pi(x)}(x)\| . \quad (2.1)$$

Now define (for this fixed x) a function of one real variable τ by

$$g(\tau) := f(p_0, \tau x + (1 - \tau)\pi(x)) - Lf_{\pi(x)}(\tau x + (1 - \tau)\pi(x)) .$$

We find that

$$\|f(p_0, x) - Lf_{\pi(x)}(x)\| = \|g(1) - g(0)\| \leq \sup\{\|g'(\tau)\| \mid 0 < \tau < 1\} .$$

However, for $\tau \in [0, 1]$,

$$g'(\tau) = [f_2(p_0, x_\tau) - f_2(p_0, \pi(x))][x - \pi(x)] ,$$

where $x_\tau := \tau x + (1 - \tau)\pi(x)$. We have by properties of the projection that $\pi(x_\tau) = \pi(x)$, so

$$\|f_2(p_0, x_\tau) - f_2(p_0, \pi(x))\| = \|f_2(p_0, x_\tau) - f_2(p_0, \pi(x_\tau))\| \leq \beta(\varepsilon),$$

since $x_\tau \in X_\varepsilon$. Hence

$$\|f(p_0, x) - Lf_{\pi(x)}(x)\| \leq \beta(\delta) \|x - \pi(x)\|. \quad (2.2)$$

As $\|f(p, x) - f(p_0, x)\| \leq \alpha_\delta(p)$, we have from (2.1) and (2.2)

$$\begin{aligned} \|Lf_{\pi(x)}(x) - f(p, x)\| &\leq \alpha_\delta(p) + \beta(\delta) \|x - \pi(x)\| \\ &< \frac{1}{2} \eta + \frac{1}{2} \eta = \eta. \end{aligned} \quad (2.3)$$

Hence, by our previous remarks $F_p(x)$ is a nonempty compact convex set for each $x \in X_\varepsilon$. Also, using (i), (ii) and (2.3) we have for $x \in X_\delta$,

$$\begin{aligned} d[F_p(x), X_0] &= d[Q(\pi(x))^{-1} [Lf_{\pi(x)}(x) - f(p, x)], Q(\pi(x))^{-1}(0)] \\ &\leq \lambda \|Lf_{\pi(x)}(x) - f(p, x)\| \\ &\leq \lambda \alpha_\delta(p) + \lambda \beta(\delta) \|x - \pi(x)\| \leq \frac{1}{2} \delta + \frac{1}{2} \delta = \delta, \end{aligned} \quad (2.4)$$

so F_p carries X_δ into itself. We have

$$\begin{aligned} \text{graph } F_p &= \{(x, y) \mid Lf_{\pi(x)}(x) - f(p, x) \in Lf_{\pi(x)}(y) + T(y)\} \\ &= \{(x, y) \mid 0 \in f(p, x) + f_2(p_0, \pi(x))(y - x) + T(y)\}. \end{aligned}$$

Using the continuity of f , f_2 , and π , together with the fact that T is closed (by maximal monotonicity), one can show without difficulty that $\text{graph } F_p$ is closed in $X_\delta \times X_\delta$. We can thus apply the Kakutani fixed-point theorem to conclude that there is some $x_p \in X_\varepsilon$ with $x_p \in F_p(x_p)$; that is,

$$Lf_{\pi(x_p)}(x_p) - f(p, x_p) \in Lf_{\pi(x_p)}(x_p) + T(x_p),$$

so

$$0 \in f(p, x_p) + T(x_p)$$

and thus $x_p \in \Sigma(p)$, which is therefore nonempty. We have

$$\text{graph } \Sigma = \{(p, x) \in U \times X_\delta \mid 0 \in f(p, x) + T(x)\};$$

this is closed in $U \times X_\delta$ by joint continuity of f and closure of T .

However, the range of Σ is contained in the compact set X_δ ; thus by [10, Lemma 4.4] Σ is actually upper semicontinuous from U to X_δ .

If $x_0 \in X_0$ then by (ii) one has $0 \in Lf_{x_0}(x_0) + T(x_0) = f(p_0, x_0) + T(x_0)$, so $x_0 \in \Sigma(p_0)$ and thus $\Sigma(p_0) \supset X_0$. On the other hand, if $x \in \Sigma(p_0)$ then $x \in X_\delta$ and $0 \in f(p_0, x) + T(x)$; therefore

$$Lf_{\pi(x)}(x) - f(p_0, x) \in Lf_{\pi(x)}(x) + T(x)$$

so that $x \in F_{p_0}(x)$. As $x \in X_\delta$, we have from (2.4) with $p = p_0$ that

$$d[x, X_0] \leq d[F_{p_0}(x), X_0] \leq \lambda \|Lf_{\pi(x)}(x) - f(p_0, x)\|.$$

But from (2.3) with $p = p_0$, we find that

$$\|Lf_{\pi(x)}(x) - f(p_0, x)\| \leq \beta(\delta) \|x - \pi(x)\| = \beta(\delta) d[x, X_0].$$

Thus

$$d[x, X_0] \leq \lambda \beta(\delta) d[x, X_0] \leq \frac{1}{2} d[x, X_0],$$

implying that $x \in X_0$ since X_0 is closed. Thus we actually have $\Sigma(p_0) = X_0$.

Now take any $\varepsilon > 0$; find $\delta_\varepsilon \in (0, \delta]$ such that for $\sigma \in [0, \delta_\varepsilon]$ one has $\lambda \beta(\sigma) \leq \frac{1}{2} \varepsilon / (\lambda + \varepsilon)$. One can show that the function

$\gamma(p) := \max\{\|f_2(p, x) - f_2(p_0, x)\| \mid x \in X_0 + \delta_\epsilon B\}$ is well defined on P and is continuous at p_0 : choose a neighborhood $U_\epsilon(p_0) \subset U$ so that if $p \in U_\epsilon$ we have $\Sigma(p) \subset \Sigma(p_0) + \delta_\epsilon B$ and $\lambda \gamma(p) \leq \frac{1}{2} \epsilon / (\lambda + \epsilon)$. Now choose any $p \in U_\epsilon$ and any $x \in \Sigma(p)$. Using (2.4) and the fact that $x \in F_p(x)$, we have

$$\begin{aligned} d[x, \Sigma(p_0)] &\leq d[F_p(x), X_0] \\ &\leq \lambda \|Lf_{\pi(x)}(x) - f(p, x)\| \\ &\leq \lambda \|h(x) - h(\pi(x))\| + \lambda \|h(\pi(x))\| \\ &\quad + \lambda \|f(p_0, x) - Lf_{\pi(x)}(x)\|, \end{aligned} \quad (2.5)$$

where $h(x) := f(p, x) - f(p_0, x)$. If we define, as before, $x_\tau := \tau x + (1 - \tau)\pi(x)$, we have

$$\|h(x) - h(\pi(x))\| \leq \|x - \pi(x)\| \sup\{\|h'(x_\tau)\| \mid 0 < \tau < 1\}.$$

But $h'(x_\tau) = f_2(p, x_\tau) - f_2(p_0, x_\tau)$, so

$$\|h(x) - h(\pi(x))\| \leq \gamma(p) \|x - \pi(x)\|.$$

Thus, using (2.2), (2.5) and the fact that $\|h(\pi(x))\| \leq \alpha_0(p)$, we have

$$\begin{aligned} d[x, \Sigma(p_0)] &\leq \lambda \gamma(p) \|x - \pi(x)\| + \lambda \alpha_0(p) + \lambda \beta(\delta_\epsilon) \|x - \pi(x)\| \\ &\leq [\epsilon / (\lambda + \epsilon)] \|x - \pi(x)\| + \lambda \alpha_0(p). \end{aligned}$$

But $\|x - \pi(x)\| = d[x, \Sigma(p_0)]$, so if $\lambda > 0$ we obtain

$$[\lambda / (\lambda + \epsilon)] d[x, \Sigma(p_0)] \leq \lambda \alpha_0(p)$$

and thus

$$d[x, \Sigma(p_0)] \leq (\lambda + \epsilon) \alpha_0(p). \quad (2.6)$$

On the other hand, if $\lambda = 0$ then (2.5) implies that $d[x, \Sigma(p_0)] = 0$, in which case (2.6) holds trivially. In either case, therefore,

$$\Sigma(p) \subset \Sigma(p_0) + (\lambda + \epsilon)\alpha_0(p)B,$$

which completes the proof.

In order to apply Theorem 1, one has to be able to verify the uniform upper Lipschitz continuity of the multifunctions $Q(x_0)^{-1}$ for $x_0 \in X_0$. As this may not always be easy to do, it is of interest to avoid this requirement by employing Theorem 2 in those cases to which it is applicable. We shall therefore proceed to establish, in the next section, results about polyhedral multifunctions which will allow us to prove Theorem 2.

3. Polyhedral multifunctions. We have already defined the class of polyhedral multifunctions, and it is clear from the definition that such a multifunction is always closed. Our goal here is to show that it is in fact locally upper Lipschitzian at each point with a uniform modulus; to do this we require several lemmas.

LEMMA 1: Let $P : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a nonempty polyhedral multifunction and let $G_i, 1 \leq i \leq k,$ be its components. Let $x_0 \in \text{dom } P,$ and define $J := \{i | x_0 \in \pi_x(G_i)\},$ where π_x is the canonical projection of $\mathbb{R}^n \times \mathbb{R}^m$ on $\mathbb{R}^n.$ Then there exists a neighborhood $U(x_0)$ such that

$$(U \times \mathbb{R}^m) \cap \text{graph } P \subset \bigcup_{i \in J} G_i.$$

PROOF: For each i , both $\{x_0\} \times \mathbb{R}^m$ and G_i are nonempty polyhedral convex sets in \mathbb{R}^{n+m} . If $i \notin J$ they do not meet, and thus

by [16, Cor. 19.3.3] they can be strongly separated. In particular, then, there is some neighborhood $U_i(x)$ such that $(U_i \times \mathbb{R}^m) \cap G_i = \emptyset$. Let $U := \bigcap_{i \notin J} U_i$; this U is a neighborhood of x_0 because the number of components is finite. Clearly,

$$(U \times \mathbb{R}^m) \cap \text{graph } P \subset \left(\bigcup_{i=1}^k G_i \right) \setminus \left(\bigcup_{i \notin J} G_i \right) \subset \bigcup_{i \in J} G_i,$$

and this completes the proof.

Lemma 1 tells us that if $x \in U$ and $y \in P(x)$, then (x, y) belongs to a component G_i which also contains (x_0, y_0) for some $y_0 \in P(x_0)$. In addition to this information, we also need an extension of a basic theorem of Hoffman [4] about linear inequalities.

LEMMA 2 (Generalized Hoffman Theorem): Let G be a nonempty polyhedral convex set in $\mathbb{R}^n \times \mathbb{R}^m$. Let π_x and π_y be the canonical projections of $\mathbb{R}^n \times \mathbb{R}^m$ onto \mathbb{R}^n and \mathbb{R}^m . For $\bar{z} := (\bar{x}, \bar{y}) \in \pi_x(G) \times \pi_y(G)$ define

$$d_x(\bar{z}, G) := \min \{ \|x - \bar{x}\| \mid (x, \bar{y}) \in G \}$$

and

$$d_y(\bar{z}, G) := \min \{ \|y - \bar{y}\| \mid (\bar{x}, y) \in G \}.$$

Then there exist ξ, η such that for each $z \in \pi_x(G) \times \pi_y(G)$

$$d_x(z, G) \leq \eta d_y(z, G); \quad d_y(z, G) \leq \xi d_x(z, G).$$

PROOF: As G is polyhedral, we can represent it in the form

$$G = \{ (x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid Ax + By \leq c \},$$

$$Lf_{\pi(x_p)}(x_p) - f(p, x_p) \in Lf_{\pi(x_p)}(x_p) + T(x_p),$$

so

$$0 \in f(p, x_p) + T(x_p)$$

and thus $x_p \in \Sigma(p)$, which is therefore nonempty. We have

$$\text{graph } \Sigma = \{(p, x) \in U \times X_\delta \mid 0 \in f(p, x) + T(x)\};$$

this is closed in $U \times X_\delta$ by joint continuity of f and closure of T .

However, the range of Σ is contained in the compact set X_δ ; thus

by [10, Lemma 4.4] Σ is actually upper semicontinuous from U to X_δ .

If $x_0 \in X_0$ then by (ii) one has $0 \in Lf_{x_0}(x_0) + T(x_0) = f(p_0, x_0) + T(x_0)$,

so $x_0 \in \Sigma(p_0)$ and thus $\Sigma(p_0) \supset X_0$. On the other hand, if $x \in \Sigma(p_0)$

then $x \in X_\delta$ and $0 \in f(p_0, x) + T(x)$; therefore

$$Lf_{\pi(x)}(x) - f(p_0, x) \in Lf_{\pi(x)}(x) + T(x)$$

so that $x \in F_{p_0}(x)$. As $x \in X_\delta$, we have from (2.4) with $p = p_0$ that

$$d[x, X_0] \leq d[F_{p_0}(x), X_0] \leq \lambda \|Lf_{\pi(x)}(x) - f(p_0, x)\|.$$

But from (2.3) with $p = p_0$, we find that

$$\|Lf_{\pi(x)}(x) - f(p_0, x)\| \leq \beta(\delta) \|x - \pi(x)\| = \beta(\delta) d[x, X_0].$$

Thus

$$d[x, X_0] \leq \lambda \beta(\delta) d[x, X_0] \leq \frac{1}{2} d[x, X_0],$$

implying that $x \in X_0$ since X_0 is closed. Thus we actually have $\Sigma(p_0) = X_0$.

Now take any $\varepsilon > 0$; find $\delta_\varepsilon \in (0, \delta]$ such that for $\sigma \in [0, \delta_\varepsilon]$

one has $\lambda \beta(\sigma) \leq \frac{1}{2} \varepsilon / (\lambda + \varepsilon)$. One can show that the function

It will be of interest to establish some simple methods for constructing polyhedral multifunctions from other such functions. To begin with, since polyhedrality was defined in terms of graphs, it is clear that the inverse of any polyhedral multifunction is itself polyhedral. It is also easy to show that if F is polyhedral and $\lambda \in \mathbb{R}$, then λF is polyhedral. The next result shows that the sum of two polyhedral multifunctions is also polyhedral.

PROPOSITION 2: Let F and H be polyhedral multifunctions from \mathbb{R}^n into \mathbb{R}^m . Then $F + H$ is polyhedral.

PROOF: If $\text{dom } F \cap \text{dom } H = \emptyset$ then $F + H$ is empty, hence certainly polyhedral. Otherwise, let G_1, \dots, G_k be the components of F and let K_1, \dots, K_ℓ be those of H . Let $\mathcal{P} := \{(i, j) \mid \pi_x(G_i) \cap \pi_x(K_j) \neq \emptyset\}$; for $(i, j) \in \mathcal{P}$ let

$$G_i \Delta K_j := \{(x, y + z) \mid (x, y) \in G_i, (x, z) \in K_j\}.$$

We note that

$$G_i \Delta K_j = C[(G_i \times K_j) \cap D(\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m)],$$

where C and D are linear transformations defined by

$$C : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n \times \mathbb{R}^m; C(x, y, v, w) := (x, y + w)$$

and

$$D : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m; D(x, y, w) := (x, y, x, w).$$

Hence $G_i \Delta K_j$ is a nonempty polyhedral convex set. Now if $(x, v) \in \text{graph}(F + H)$ then $v = y + z$ with $(x, y) \in \text{graph } F$ and $(x, z) \in \text{graph } H$.

For some i and j , we have $(x, y) \in G_i$, $(x, z) \in K_j$, and $(i, j) \in P$, so that $(x, y + z) \in G_i \Delta K_j$ and hence

$$\text{graph}(F + H) \subset \bigcup_{(i, j) \in P} G_i \Delta K_j.$$

On the other hand, if (x, v) belongs to the set on the right then for some $(i, j) \in P$, $(x, v) \in G_i \Delta K_j$. Hence $v = y + z$ with $(x, y) \in G_i \subset \text{graph } F$ and $(x, z) \in K_j \subset \text{graph } H$, so we have $(x, v) \in \text{graph}(F + H)$. Thus

$$\text{graph}(F + H) = \bigcup_{(i, j) \in P} G_i \Delta K_j,$$

and therefore $F + H$ is polyhedral, as was to be shown.

We can now establish Theorem 2, using these results.

PROOF OF THEOREM 2: We shall prove that the hypotheses of Theorem 1 are satisfied; indeed, those listed under (ii) and (iii) are certainly satisfied under the assumptions of Theorem 2. To prove (i), it suffices to show that the multifunction $[A(\cdot) + a + T(\cdot)]^{-1}$ is locally upper Lipschitzian at the origin. It is evident that $A(\cdot) + a$ is polyhedral, whereas T is polyhedral by assumption; hence by Proposition 2 $A(\cdot) + a + T(\cdot)$ is polyhedral, and thus so is its inverse. The upper Lipschitzian property now follows from Proposition 1, and this proves Theorem 2.

We conclude this section by proving our earlier assertion that if f is a polyhedral convex function which never takes the value $-\infty$, then ∂f is a polyhedral multifunction. The proof employs two preliminary lemmas, in which for $y, z \in \mathbb{R}^n$ we use the notation $[y, z]$ to mean $\{(1 - \lambda)y + \lambda z \mid 0 \leq \lambda \leq 1\}$.

LEMMA 3: If C is nonempty, closed and convex in \mathbb{R}^n with $y, z \in C$, then for each $x \in \text{ri}[y, z]$, $\partial\psi_C(x) \subset \partial\psi_C(y)$. Consequently, $\partial\psi_C(\cdot)$ is constant on the relative interior of any convex subset of C .

PROOF: As $x \in \text{ri}[y, z]$ we can write $x = (1 - \lambda)y + \lambda z$, with $\lambda \in (0, 1)$. Suppose $x^* \in \partial\psi_C(x)$ (which is nonempty since $x \in C$); let c be any point in C . We have $(1 - \lambda)c + \lambda z \in C$; thus

$$\langle x^*, [(1 - \lambda)c + \lambda z] - x \rangle \leq 0.$$

But $[(1 - \lambda)c + \lambda z] - x = (1 - \lambda)[c - (1 - \lambda)^{-1}(x - \lambda z)] = (1 - \lambda)(c - y)$; consequently, $\langle x^*, c - y \rangle \leq 0$ so $x^* \in \partial\psi_C(y)$. Thus $\partial\psi_C(x) \subset \partial\psi_C(y)$.

To prove the second assertion, let F be a convex subset of C . If F is empty or a singleton the result is immediate; therefore suppose that F (and hence $\text{ri } F$) contains more than one point: let $x, y \in \text{ri } F$ with $x \neq y$. By [16, Th. 6.1], there exist $\hat{x}, \hat{y} \in F$ and $\lambda, \mu \in (0, 1)$ such that

$$x = (1 - \lambda)\hat{x} + \lambda y, \quad y = (1 - \mu)\hat{y} + \mu x.$$

But then $x \in \text{ri}[\hat{x}, y]$, so $\partial\psi_C(x) \subset \partial\psi_C(y)$, and $y \in \text{ri}[\hat{y}, x]$, so $\partial\psi_C(y) \subset \partial\psi_C(x)$. Thus $\partial\psi_C(x) = \partial\psi_C(y)$, and so $\partial\psi_C(\cdot)$ is constant on $\text{ri } F$, which completes the proof.

We can now prove that $\partial\psi_C$ is polyhedral for polyhedral C .

LEMMA 4: Let C be a nonempty polyhedral convex set in \mathbb{R}^n .

Then $\partial\psi_C$ is a polyhedral multifunction.

PROOF: By [16, Th. 19.1], C has finitely many faces F_1, \dots, F_m , and by [16, Th. 18.2] the collection $\{\text{ri } F_i\}_{i=1}^m$ is a partition of C . By

Lemma 3, on each set $\text{ri } F_i$ the function $\partial\psi_C(\cdot)$ has a constant value, a set which we will call P_i . By [16, Th. 23.10], P_i is a nonempty polyhedral convex set. Thus,

$$\text{graph } \partial\psi_C = \bigcup_{i=1}^m [(\text{ri } F_i) \times P_i] .$$

Now consider the set

$$\Gamma := \bigcup_{i=1}^m (F_i \times P_i) .$$

Clearly $\text{graph } \partial\psi_C \subset \Gamma$. On the other hand, suppose that $(x, x^*) \in \Gamma$; then for some i , $x \in F_i$ and $x^* \in P_i$. If $x \in \text{ri } F_i$ then $(x, x^*) \in \text{graph } \partial\psi_C$. If $x \notin \text{ri } F_i$, then there is some $z \neq x$ with $z \in \text{ri } F_i$ (since $\text{ri } F_i \neq \emptyset$). Using [16, Th. 6.1] we can find $y \in F_i$ and $\lambda \in (0, 1)$ such that $z = (1 - \lambda)y + \lambda x$. But then $z \in \text{ri}[y, x]$, and by Lemma 3 we have $P_i = \partial\psi_C(z) \subset \partial\psi_C(x)$. However, as $\{\text{ri } F_i\}_{i=1}^m$ is a partition of C and as $x \in C$ (since $F_i \subset C$ for each i), there exists some j such that $x \in \text{ri } F_j$. But then $\partial\psi_C(x) = P_j$ and we have $P_i \subset P_j$. As $x^* \in P_i$, we have $(x, x^*) \in (\text{ri } F_j) \times P_j \subset \text{graph } \partial\psi_C$. Hence

$$\text{graph } \partial\psi_C = \Gamma = \bigcup_{i=1}^m (F_i \times P_i) .$$

But each F_i is a polyhedral convex set (as a face of C), so $F_i \times P_i$ is polyhedral and $\partial\psi_C$ is a polyhedral multifunction. This completes the proof of Lemma 4.

PROPOSITION 3: Let f be a polyhedral convex function from \mathbb{R}^n into $(-\infty, +\infty]$. Then ∂f is a polyhedral multifunction.

PROOF: If $f \equiv +\infty$ then $\partial f \equiv \mathbb{R}^n$ so we have the result. We may therefore assume that f is proper, so that its epigraph \mathcal{E} is non-empty. We shall first prove that for any proper convex function f , if D is the linear transformation from $\mathbb{R}^1 \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$ to $\mathbb{R}^n \times \mathbb{R}^n$ which takes (x, ξ, y, η) into (x, y) (i.e., projects onto the first and third of the spaces involved) then

$$\text{graph } \partial f = D[(\text{graph } \partial \psi_{\mathcal{E}}) \cap (\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \{-1\})] . \quad (3.1)$$

For the proof, let $(x, x^*) \in \text{graph } \partial f$. Then $f(x) < +\infty$, and $(x, x^*) = D(x, f(x), x^*, -1)$. Let $(y, \eta) \in \mathcal{E}$; then $\eta \geq f(y) \geq f(x) + \langle x^*, y - x \rangle$, so that

$$\langle (x^*, -1), (y, \eta) - (x, f(x)) \rangle \leq 0 ,$$

and hence $(x^*, -1) \in \partial \psi_{\mathcal{E}}(x, f(x))$, so

$$(x, f(x), x^*, -1) \in \text{graph } \partial \psi_{\mathcal{E}} .$$

This establishes the inclusion \subset ; for the other direction let

$(x, \xi, x^*, -1) \in \text{graph } \partial \psi_{\mathcal{E}}$; we will show that $(x, x^*) \in \text{graph } \partial f$. Consider any $y \in \mathbb{R}^n$. If $f(y) < +\infty$ then $(y, f(y)) \in \mathcal{E}$, so that

$$\langle (x^*, -1), (y, f(y)) - (x, \xi) \rangle \leq 0 .$$

Thus, as $f(x) \leq \xi$,

$$\langle x^*, y - x \rangle \leq f(y) - \xi \leq f(y) - f(x) ,$$

so that

$$f(y) \geq f(x) + \langle x^*, y - x \rangle .$$

But if $f(y) = +\infty$ then the last inequality holds trivially, and so

$(x, x^*) \in \text{graph } \partial f$. This establishes (3.1).

Now if f is polyhedral then by definition \mathcal{E} is a polyhedral convex set; hence by Lemma 4 $\text{graph } \partial \psi_{\mathcal{E}}$ is the union of finitely many polyhedral convex sets, say P_1, \dots, P_m . Denote by F the set $\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \{-1\}$, also obviously polyhedral convex. Using (3.1), we have

$$\begin{aligned} \text{graph } \partial f &= D\left[\left(\bigcup_{i=1}^m P_i\right) \cap F\right] \\ &= D\left[\bigcup_{i=1}^m (P_i \cap F)\right] \\ &= \bigcup_{i=1}^m [D(P_i \cap F)], \end{aligned}$$

and since $D(P_i \cap F)$ is polyhedral for each i , we see that ∂f is a polyhedral multifunction. This proves Proposition 3.

4. An application: Newton's method for variational inequalities.

We shall analyze here an algorithm for solving variational inequalities over polyhedral convex sets. To specify precisely the form of the variational inequalities with which we shall deal, we take x_* to be a point in \mathbb{R}^n , U a neighborhood of x_* , and f a (Fréchet) differentiable function from U into \mathbb{R}^{n+m} . We let A be a linear transformation from \mathbb{R}^m into \mathbb{R}^{n+m} , and take a to be a point in \mathbb{R}^{n+m} and C a nonempty polyhedral convex set in \mathbb{R}^{n+m} . We shall consider the variational inequality

$$0 \in f(x) + Ay + a + \partial\psi_C(x, y) . \quad (4.1)$$

Of course, we could have $m = 0$, so that A does not appear (we could, alternatively, have $n = 0$, but then the algorithm we will discuss is of no interest as the problem is already linear). To solve (4.1), we propose to use an algorithm which generalizes the well-known Newton method for solving systems of nonlinear equations [11]. In that method, one proceeds by selecting a trial solution, linearizing the nonlinear function whose zero one is trying to find, then solving the resulting system. Thus, one retains the form of the system (equations) but substitutes the linearized function for the original nonlinear one.

The algorithm we propose is based upon the point of view that variational inequalities generalize equations (which correspond to the case in which C is the whole space). Thus, in (4.1) we replace $f(x)$ by its linearization

$$Lf_{x_k}(x) := f(x_k) + f'(x_k)(x - x_k)$$

about a point $x_k \in U$, and try to solve the linear variational inequality

$$0 \in Lf_{x_k}(x) + Ay + a + \partial\psi_C(x, y) , \quad (4.2)$$

thereby retaining the original form of the problem as a variational inequality but replacing the nonlinear function f by its linearization about x_k .

This prescription raises obvious questions, such as whether the sub-problems (4.2) will even be solvable and, if they are, whether the

solutions will converge, to what, and how quickly. We shall prove here that if the initial point x_0 is chosen close enough to a solution of (4.1) having certain desirable properties, and if the succeeding points are chosen from a certain neighborhood, then the sub-problems (4.2) will be solvable and the sequence of points will converge quadratically to the solution set of (4.1). The result is stated precisely in the next theorem, in which we use the notation

$$S(w) := \{(x, y) \in \mathbb{R}^{m+n} \mid 0 \in Lf_w(x) + Ay + a + \partial\psi_C(x, y)\}.$$

THEOREM 3: Let x_* , U , C , f , A and a be as defined above.
Suppose that f' is Lipschitzian with modulus L on U , that $[f'(x_*) A]$
is positive semidefinite, and that for some nonempty compact convex set
 $Y_* \subset \mathbb{R}^m$, $S(x_*) = \{x_*\} \times Y_*$.

Then for any $\epsilon > 0$ there exist $\delta > 0$ and $\rho > 0$ such that for
each $x \in x_* + \rho B$ one has, with $S_\delta := S(x_*) + \delta B$,

$$\emptyset \neq S(x) \cap S_\delta \subset S(x_*) + \frac{1}{2}(\lambda + \epsilon)L\|x - x_*\|^2 B,$$

where λ is a local upper Lipschitz modulus for the polyhedral multifunction

$$M(z) := \{(x, y) \mid z \in Lf_{x_*}(x) + Ay + a + \partial\psi_C(x, y)\}.$$

We remark that λ must exist (by Proposition 1).

PROOF: We shall apply Theorem 2 with $P := U$, $N := \mathbb{R}^n \times \mathbb{R}^m$,
 $p_0 := x_*$, $X_0 := \{x_*\} \times Y_* = S(x_*)$, $T := \partial\psi_C$, and
 $f(p, (x, y)) := Lf_p(x) + Ay + a.$

We note that if $y_* \in Y_*$ then $f_2(p_0, (x_*, y_*)) = [f'(x_*) A]$, which is positive semidefinite by hypothesis; also

$$\begin{aligned} & f(p_0, (x_*, y_*)) + f_2(p_0, (x_*, y_*))[(x, y) - (x_*, y_*)] \\ &= [f'(x_*) A](x, y) + [f(x_*) - f'(x_*)x_* + a], \end{aligned}$$

so that the constancy assumption (b) of Theorem 2 holds, whereas assumption (a) holds by hypothesis. Applying Theorem 2, we conclude that there exist $\delta > 0$ and $\rho > 0$ such that for any $x' \in x_* + \rho B$, the set

$$\Sigma(x') := \{(x, y) \in S(x_*) + \delta B \mid 0 \in Lf_{x'}(x) + Ay + a + \partial\psi_C(x, y)\}$$

is nonempty and is contained in $S(x_*) + (\lambda + \epsilon)\alpha_0(x')B$. However,

$$\begin{aligned} \alpha_0(x') &= \max\{\|(Lf_{x'}(x) + Ay + a) - (Lf_{x_*}(x) + Ay + a)\| \mid (x, y) \in X_0\} \\ &= \|Lf_{x'}(x_*) - Lf_{x_*}(x_*)\| = \|f(x_*) - f(x') - f'(x')(x_* - x')\|, \end{aligned}$$

and a standard bounding technique using the Lipschitz continuity of f' shows that this quantity is bounded above by $\frac{1}{2}L\|x' - x_*\|^2$. Thus

$$\emptyset \neq S(x') \cap S_\delta \subset S(x_*) + \frac{1}{2}(\lambda + \epsilon)L\|x' - x_*\|^2 B,$$

which completes the proof of Theorem 3.

Theorem 3 shows that if we choose x_0 with $\|x_0 - x_*\| \leq \rho$ and $\sigma := \frac{1}{2}(\lambda + \epsilon)L\|x_0 - x_*\| < 1$, then select (x_1, y_1) from $S(x_0) \cap S_\delta$, we have

$$d[(x_1, y_1), S(x_*)] \leq \frac{1}{2}(\lambda + \epsilon)L\|x_0 - x_*\|^2 \leq \sigma\|x_0 - x_*\| \leq \rho.$$

If $L = 0$ or $x_0 = x_*$ we are finished; otherwise, we have

$$\frac{1}{2}(\lambda + \epsilon)L\|x_1 - x_*\| \leq \left[\frac{1}{2}(\lambda + \epsilon)L\|x_0 - x_*\|\right]^2 = \sigma^2 < \sigma,$$

so that x_1 satisfies the same assumptions as did x_0 . Accordingly $S(x_1)$ is nonempty, and selecting (x_{k+1}, y_{k+1}) from $S(x_k) \cap S_\delta$ for $k = 1, 2, \dots$ we find by an easy induction that the sequence $\{(x_{k+1}, y_{k+1})\}$ exists and satisfies

$$d[(x_{k+1}, y_{k+1}), S(x_*)] \leq \frac{1}{2}(\lambda + \varepsilon)L\|x_k - x_*\|^2 \leq 2(\lambda + \varepsilon)^{-1}L^{-1}_\sigma(2^{k+1}).$$

Hence the sequence $\{d[(x_k, y_k), S(x_*)]\}$ converges to zero, and the convergence is R-quadratic in the sense of [11].

The requirement that (x_{k+1}, y_{k+1}) be chosen from S_δ may be difficult or impossible to implement computationally. It can, however, be avoided by placing more assumptions on the functions appearing in the problem. For example, if f and A are such that one knows that the solution set of each sub-problem (4.2) is connected, then S_δ can be eliminated since, in Theorem 1, the upper semicontinuity of Σ and the fact that $\Sigma(p_0) = X_0$ imply that for p near p_0 the solution set of $0 \in f(p, x) + \partial\psi_C(x)$ either lies entirely in X_δ or is disconnected.

At this point, the algorithm we have presented has to be regarded as an exercise in constructive real analysis rather than as an effective computational device, since large-scale testing of its effectiveness on real problems has not yet been carried out. However, in one special case some experience is available: for the nonlinear programming problem (1.5) the method proposed here reduces to that of Wilson [20], for which some computational results were reported in [1]. These were comparatively

unpromising, perhaps because Wilson's code used an early simplex algorithm for solving the quadratic programming problems to which the sub-problems (4.2) reduce in the case he considered. Later computational investigation of a method similar to Wilson's produced much more promising solution times [12], and analytic investigation proved the convergence to be quadratic under certain regularity assumptions [13].

Actually, Wilson proposed in a manuscript [21] a method for solving the following problem, which extends the usual nonlinear complementarity problem:

$$H(x, y) = 0$$

$$x, y \geq 0$$

$$\langle x, y \rangle = 0$$

(to recover the nonlinear complementarity problem, take $H(x, y) := F(x) - y$).

His idea was to partition the variables into basic and nonbasic sets, then solve a system of linear equations obtained by linearizing H with respect to the basic variables while holding the nonbasic variables constant at zero. In case this procedure resulted in negative basic variables, he proposed to use a pivoting strategy to eliminate the negative variable from the basis. His procedure seems very close to what would be done by the algorithm proposed above when applied to a nonlinear complementarity problem; however, the author is unaware of any formal analysis of the convergence and rate of convergence of Wilson's proposed method.

Finally, Kuhn [6] reported work of Saigal (privately communicated) on the use of vector labelings in pivotal algorithms for computing fixed points; according to Kuhn, Saigal had shown how the use of such techniques could lead to quadratic approximations in simplicial-subdivision techniques. This work may be related to the algorithm we present here.

Added note. The work of Saigal, mentioned above, appears in a manuscript entitled, "On the convergence rate of algorithms for solving equations that are based on methods of complementary pivoting," a copy of which was kindly furnished to the author by Dr. Saigal. It deals with the problem of finding a fixed point of a Lipschitz continuously differentiable function $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$. Thus, it appears to be directly related to the work presented here only for the case in which $T(x)$ is identically zero: i.e., the case of conventional systems of nonlinear equations, for which our algorithm reduces to the standard Newton method.

REFERENCES

1. A. R. Colville, "A Comparative Study on Nonlinear Programming Codes," IBM New York Scientific Center Report No. 320-2949, 1968.
2. R. W. Cottle, "Nonlinear Programs with Positively Bounded Jacobians," SIAM J. Appl. Math. 14, 1966, pp. 147-158.
3. G. B. Dantzig and R. W. Cottle, "Positive (Semi-) Definite Programming," in: Nonlinear Programming, ed. J. Abadie, North-Holland, Amsterdam, 1968, pp. 55-73.
4. A. J. Hoffman, "On Approximate Solutions of Systems of Linear Inequalities," J. Res. Nat. Bur. Standards 49, 1952, pp. 263-265.
5. W. W. Hogan, "Project Independence Evaluating System Integrating Model," Working Paper CC-74-71, Office of Quantitative Methods, Federal Energy Administration, Washington, D. C., 1974.
6. H. W. Kuhn, "How to Compute Economic Equilibria by Pivotal Methods," to appear in: Computing Equilibria: How and Why, eds. J. Łoś and M. W. Łoś, North-Holland, Amsterdam, 1976.
7. O. L. Mangasarian, Nonlinear Programming, McGraw-Hill, New York, 1969.
8. J. J. More, "Coercivity Conditions in Nonlinear Complementarity Problems," SIAM Review 16, 1974, pp. 1-16.
9. J. J. More, "Classes of Functions and Feasibility Conditions in Nonlinear Complementarity Problems," Math. Programming 6, 1974, pp. 327-338.

10. H. Nikaido, Convex Structures and Economic Theory, Academic Press, New York and London, 1968.
11. J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York and London, 1970.
12. S. M. Robinson, "A Quadratically-Convergent Algorithm for General Nonlinear Programming Problems," Math. Programming 3, 1972, pp. 145-156.
13. S. M. Robinson, "Perturbed Kuhn-Tucker Points and Rates of Convergence for a Class of Nonlinear-Programming Algorithms," Math. Programming 7, 1974, pp. 1-16.
14. R. T. Rockafellar, "Local Boundedness of Nonlinear, Monotone Operators," Michigan Math. J. 16, 1969, pp. 397-407.
15. R. T. Rockafellar, "On the Maximality of Sums of Nonlinear Monotone Operators," Trans. Amer. Math. Soc. 149, 1970, pp. 75-88.
16. R. T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, NJ, 1970.
17. R. T. Rockafellar, "Monotone Operators and the Proximal Point Algorithm," to appear in SIAM J. Control and Optimization, 1976.
18. H. Scarf, The Computation of Economic Equilibria, Yale University Press, New Haven and London, 1973.
19. M. H. Wagner, "Project Independence Evaluation System Integrating Model," manuscript.

20. R. B. Wilson, A Simplicial Algorithm for Concave Programming,
Dissertation, Graduate School of Business Administration, Harvard
University, Boston, 1963.
21. R. B. Wilson, "Subroutine SOLVER Description," mimeographed
manuscript, 1964.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 1672 (14) MRC-TSR-1672	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER (9) Technical
4. TITLE (and Subtitle) AN IMPLICIT-FUNCTION THEOREM FOR GENERALIZED VARIATIONAL INEQUALITIES	5. TYPE OF REPORT & PERIOD COVERED Summary Report, no specific reporting period	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Stephen M. Robinson	8. CONTRACT OR GRANT NUMBER(s) DAAG29-75-C-0024 NSF-DCR74-20584	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of 610 Walnut Street Wisconsin Madison, Wisconsin 53706	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (12) 34P.	
11. CONTROLLING OFFICE NAME AND ADDRESS See Item 18 below.	12. REPORT DATE (11) Sep 1976	13. NUMBER OF PAGES 31
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) UNCLASSIFIED	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES U. S. Army Research Office P. O. Box 12211 Research Triangle Park North Carolina 27709 National Science Foundation Washington, D. C. 20550		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Variational inequalities Monotone operators Nonlinear complementarity problem Nonlinear programming Economic equilibria		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) An implicit-function theorem is established for a multifunction consisting of the sum of a differentiable function and a maximal monotone operator. Applications to nonlinear complementarity problems, mathematical programming problems, and economic equilibria are pointed out. An application to the analysis of a general Newton method for solving variational inequalities is treated in some detail.		